Clinics of Oncology

Prognostic Value of LINC01600 and CASC15 as Competitive Endogenous RNAs in Lung Adenocarcinoma

Zhou F[1*]

[1]Department of Tropical Diseases, Naval Medical University, Shanghai, China

*Corresponding author:
Fangbin Zhou,
Department of Tropical Diseases, Naval
Medical University, Shanghai, China,
E-mail: 18801768978@139.com

Citation:

Zhou F. Prognostic Value of LINC01600 and CASC15 as Competitive Endogenous RNAs in Lung Adenocarcinoma. Clin Onco. 2022; 6(13): 1-12

Abbreviations:

lncRNAs: long non-coding RNAs; miRNAs: microRNAs; LUAD: lung adenocarcinoma; HGNC: The HUGO Gene Nomenclature Committee; TCGA: The Cancer Genome Atlas; TKIs tyrosine kinase inhibitors; ncRNAs, non-coding RNA; MREs: microRNA response elements; ceRNAs: competing endogenous RNAs; PTEN: tensin homologue; GEO: Gene Expression Omnibus; CGP: the Wellcome Trust Sanger Institute's Cancer Genome Project; GO: gene ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; PPI DAVID: Database for Annotation, Visualization and Integrated Discovery; BP: biological process; MF: molecular function; CC: cellular component

1. Abstract

Long non-coding RNAs (lncRNAs) can directly or indirectly regulate gene expression through interacting with microRNAs (miRNAs). Competitive Endogenous RNAs (ceRNAs) render the roles of lncRNAs more complicated in the process of tumor occurrence and progression. However, the prognostic value of lncRNAs as potential biomarkers and their functional roles as ceRNAs have not been clearly described for lung adenocarcinoma (LUAC). In the present study, the aberrant expression profiles of lncRNAs and miRNAs were analyzed at cBioPortal by interrogating LUAD dataset from The Cancer Genome Atlas (TCGA) database with 517 tissue samples. A total of 92 lncRNAs and 125 miRNAs with highly genetic alterations were identified. Further bioinformatics analysis was performed to construct a LUAD-related lncRNA-miRNA-mRNA ceRNA network, which included 24 highly altered lncRNAs, 21 miRNAs and 142 mRNAs. Some key lncRNAs in this network were subsequently identified as LUAD prognosis-related, and of those, LINC01600 and CASC15 both performed the potential prognostic characteristics with LUAD regarding OS and recurrence. Comprehensive analysis indicated that the expression of LINC01600 was significantly associated with KRAS mutation and lymph node metastasis, and CASC15 and LINC01600 were significantly tended towards co-occurrence, which may be due to the similarity of genes co-expressed by these two lncRNAs. Our findings provided novel insight into better understanding of ceRNA regulatory mechanisms in the pathogenesis of LUAD and facilitated the identification of potential biomarkers for prognosis.

2. Introduction

Lung cancer is the most frequently reported malignancy worldwide, leading to more than 2.21 million incidences and almost 1.80 million deaths, as estimated in 2020 [1]. Lung adenocarcinoma (LUAD) is the most common histological subtype of lung cancer, which accounts for an approximate proportion of 40%. There is a huge of pivotal molecules like TKIs (tyrosine kinase inhibitors) identified for targeted therapies, which have dramatically improved treatment for LUAD patients [2-5]. However, the overall 5-year survival rate of LUAD patients has not been significantly improved over the past few decades. Therefore, it is urgently required to understand the molecular nature of LUAD pathogenesis and develop new targets for the diagnosis and treatment of LUAD.

Increasing evidence has emphasized the involvement of non-coding RNAs (ncRNAs) in the process of tumor occurrence and progression [6,7]. ncRNAs are a class of RNA molecules without protein-coding function, and microRNAs (miRNAs) and long

non-coding RNAs (lncRNAs) are two typical subtypes. miRNAs can repress the expression of targeted mRNAs by partial complementation binding to the mRNA sequences, which named microRNA response elements (MREs) [8]. It has been well-documented that miRNA-mRNA interactions involve a variety of biological processes, including carcinogenesis and tumor metastasis [9,10]. In contrast, lncRNAs function via more multiple mechanisms [11]. In addition to directly targeting mRNAs, lncRNAs can function as competing endogenous RNAs (ceRNAs) to indirectly regulate mRNAs through competing for shared miRNAs. In 2011, ceRNA hypothesis was proposed by Salmena and colleagues, which states that lncRNAs, mRNAs and other RNAs can act as natural miRNA sponges to inhibit miRNA function by competitive binding to shared MREs [12]. Later, this hypothesis was updated and validated experimentally by more and more convincing evidences [13-15]. Pandolfi et al. demonstrated the existence of phosphatase and tensin homologue (PTEN) and its ceRNAs interaction among mRNAs in vitro and in vivo [16]. Kallen et al. found that H19 can affect the expression of endogenous let-7 targets through acting as a molecular sponge to sequester let-7 [17]. The above investigations regarding the lncRNA-miRNA-mRNA ceRNA networks provide a better understanding of the roles of lncRNA-miRNA interactions in mRNAs regulation and tumorigenesis.

With the development of multi-centered cancer genomic projects, such as The Cancer Genome Atlas (TCGA), Gene Expression Omnibus (GEO) and the Wellcome Trust Sanger Institute's Cancer Genome Project (CGP), bioinformatic technologies and tools for high throughput genome analysis were increasingly created and more and more researches focused on mining database to study ceRNA networks in cancer. Katarzyna summarized a highly illustrated review of the common seven web-based portals for the interpretation of the data. All of the selected portals were user-friendly and expected to contribute to a better understanding of cancer molecular etiology and the translation of genomic knowledge into clinical practice. In addition, to annotate the associations between ncRNAs and cancers, some disease-associated algorithms and databases, such as StarBase, miRTarBase, and DIANA-LncBase/TarBase were developed [18-21]. The integration of web portals and datasets made it feasible to construct a ceRNA regulatory network in cancer. Recently, based on TCGA or GEO data analysis, some lncRNA-miRNA-mRNA ceRNA regulatory networks were constructed among gastric, breast, and bladder cancer, and numerous prognosis-related ncRNA biomarkers were identified [22-24]. A few studies also used different sample datasets to investigate the LUAD related lncRNA-miRNA-mRNA ceRNA networks and identify several prognostic lncRNAs [25,26]. However, comprehensive analysis of ncRNAs related to LUAD at genome-wide RNA profiles and large sample scale were limited and the relationship between cancer specifc lncRNAs and clinical features still remained to be elucidated.

In this study, we interrogated the LUAD dataset of TCGA containing 517 samples at cBioPortal with the latest updated lncRNAs and miRNAs database from The HUGO Gene Nomenclature Committee (HGNC) and constructed a regulated lncRNA-miRNA-mRNA based ceRNA networks. A total of 24 lncRNAs, 19 miRNAs and 142 mRNAs were involved in this network and several lncRNAs were identified as LUAD prognosis-related. Furthermore, CASC15 and LINC01600 were proven to perform the potential prognostic characteristics with LUAD regarding both OS and recurrence and were further disclosed the correlation with the clinical features and mutual exclusivity status.

3. Materials and Methods

3.1. cBioPortal and LUAD TCGA dataset

The cBioPortal for Cancer Genomics (http://www.cbioportal.org) is an open-access resource, which provides visualization, analysis tools for more than 5,000 tumor samples from 105 cancer studies in TCGA pipeline. The customized search interface enabled researchers to interactively and individually explore genetic alterations across samples from other cancer studies and specific genes. A total of 586 LUAD tissue samples were retrieved from Lung Adenocarcinoma (TCGA, Provisional) dataset. In this study, Samples without completed data (including age, gender, race, TNM stage, and histological subtype and grade) for analysis or not histologically diagnosed as LUAD were excluded and a total of 517 LUAD tissue samples were obtained. This study is in accordance with the publication guidelines provided by TCGA (https://cancergenome.nih.gov/publications/publicationguidelines). As the data comes from the TCGA database, further approval by the Ethics Committee was not required

3.2. Identification of lncRNAs and miRNAs with genetic alterations

Although there are increasing numbers of human lncRNAs and miRNAs reported from databases, the nomenclature of lncRNAs and miRNAs is still disordered. In this study, a total of 3885 lncRNAs and 1878 miRNAs were downloaded from HUGO gene nomenclature committee (HGNC) (http://www.genenames.org/) for our analysis (Supplementary Table 1 and 2). All lncRNAs and mRNAs were input and recognized in the cBioPortal, and those of which were not included in this database were excluded. Alterations in the somatic mutations, DNA copy number alterations (CNAs) from GISTIC, mRNA and microRNA expression were analyzed and calculated based on the clinical data from cBioPortal.

3.3. lncRNA-miRNA interaction analysis

The miRNAs targeted by highly altered lncRNAs were predicted using starBase (version 2.0, http://starbase.sysu.edu.cn/) and DIANA-LncBase (Experimental v.2, http://carolina.imis.athena-innovation.gr/diana_tools/web/index.php?r=lncbasev2%2Findex-experimental) databases. Results from these two databases were combined and intersected with the above highly altered

miRNAs. The intersection contained highly altered miRNAs were targeted by highly altered lncRNAs. At last, several LUAD-related lncRNA-miRNA pairs were thus obtained.

3.4. Prediction of mRNAs targeted by miRNAs and construction of a LUADD-related lncRNA-miRNA-mRNA ceRNA network

mRNAs targeted by miRNAs were retrieved from miRTarBase (version 6.0, http://mirtarbase.mbc.nctu.edu.tw) and DIANA- TarBase (version 8.0, http://carolina.imis.athena-innovation.gr/diana_tools/web/index.php?r=tarbasev8%2Findex). miRTarBase is a database that has accumulated more than three hundred and sixty thousand miRNA-target interactions (MTIs), which are collected by manually surveying pertinent literature and mining the text systematically to filter research articles related to functional studies of miRNAs. DIANA-TarBase was a database with experimentally supported miRNA-mRNA interations. In this part, mRNA targets were predicted via experimental module and had a high Pr. score. miRNAs with less than 5% genetic alteration screened via cBiPortal were excluded, and those remaining miRNAs targeted by highly altered miRNAs were extracted and then integrated with differentially expressed miRNA-mRNA regulatory relationships, generating a LUAD-related miRNA-mRNA regulation network.

The above lncRNA-miRNA and miRNA-mRNA regulatory networks were combined to obtain a comprehensive lncRNAmiRNA-mRNA ceRNA regulatory network. The network graph was constructed and visualized using Cytoscape v3.0 (http://www.cytoscape.org/). Cytoscape is an open-source software platform for visualizing molecular interaction networks and biological pathways and integrating these networks with annotations, gene expression profiles and other state data.

3.5. Functional and pathway annotation of mRNAs in the ceRNA network

In order to reveal LUAD-related biological functions and pathways, genes involved in ceRNA regulatory network were input into the Database for Annotation, Visualization and Integrated Discovery (DAVID) for GO (gene ontology) biological process analysis and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis. Fisher's exact test was used during the enrichment process. At last, STRING software was used to visually map the PPI network of targeted genes.

3.6. Identification of prognosis-associated lncRNAs

The expression value of each lncRNAs in the ceRNA network and the survival information of each sample were extracted from TCGA dataset. The cBioPortal generated the Kaplan-Meier curve following input of the gene name for disease or progression-free-survival. P<0.05 was considered to indicate a statistically significant difference.

4. Results

4.1. Genetic alteration profiles of lncRNAs and miRNAs in LUAD

To draw the genetic alteration profiles of lncRNAs in LUAD, A total of 3855 lncRNAs downloaded from HGNC were selected and performed a primary search in cBioPortal (Supplementary Table 1). The results indicated that the profile of lncRNAs genetic alterations was heterogeneous and divided into three groups based on the altered frequencies (Supplementary Table 3). The first, and smallest group, which comprises roughly 0.21% (8/3885) of the lncRNAs database, includes lncRNAs that are highly variable and had alterations in 10% cases or above. A second, larger group includes additional, frequently variable lncRNAs that had alterations more than 5% but less than 10% cases (2.19%, 85/3885). The above 92 lncRNAs defined highly altered lncRNAs in this study (Supplementary Table 3). The third group contains lncRNAs that were rarely variable (less than 5%) or no alterations (97.61%, 3792/3885).

The genetic alteration profile of 1878 miRNAs was also divided into three groups as Supplementary Table 2 shown. miRNAs that had alterations in 10% cases or above accounted for 1.28% (24/1878) and alterations more than 5% but less than 10% cases accounted for 5.38% (101/1878). The above 125 lncRNAs defined highly altered miRNAs (Supplementary Table 4). The remaining 93.34% (1753/1878) stood for the rare cases (less than 5%) or no alterations.

4.2. LncRNA-miRNA interaction network predicted by StarBase and DIANA-LncBase

Although non-coding RNAs play vitally physiological roles in the initiation and progression of LUAD, the precise mechanism and the regulation network of lncRNAs and miRNAs remain to be elucidated. Since miRNAs are interacting with lncRNAs through MREs within ceRNA network, the first thing we should do was to screen the potential MREs in lncRNAs. The database of miRNA-lncRNA interactions was downloaded from StarBase, which was performed to screen the potential MREs. 125 highly altered miRNAs were selected to be interrogated whether these miRNAs are targeted on 92 highly altered lncRNAs. The results showed that 36 of 92 lncRNAs were included in this database and 21 of 125 miRNAs may interact with 24 of 36 lncRNAs. Finally, a total of 54 lncRNA-miRNA pairs were obtained (Table 1).

Table 1: Putative miRNAs targeting lncRNA

lncRNAs	miRNAs
CASC15	hsa-mir-298
DANCR	hsa-mir-135b
DLEU2	hsa-mir-30b, hsa-mir-30d
HCG18	hsa-mir-214, hsa-mir-296, hsa-mir-148a, hsa-mir-320a, hsa-mir-30b, hsa-mir-30d
HCP5	hsa-mir-214, hsa-mir-29c
HOTAIR	hsa-mir-214, hsa-mir-148a
JPX	hsa-mir-708
LINC00174	hsa-mir-148a, hsa-mir-31
LINC00242	hsa-mir-214
LINC00339	hsa-mir-214, hsa-mir-148a
LINC00467	hsa-mir-4735
LINC00652	hsa-mir-214
LINC00662	hsa-mir-190b
LINC00667	hsa-mir-148a, hsa-mir-31
LINC01600	hsa-mir-298
LINC00696	hsa-miR-1273a, hsa-miR-4756
MALAT1	hsa-mir-92b, hsa-mir-135b, hsa-mir-708, hsa-mir-383, hsa-mir-320a, hsa-mir-30b, hsa-mir-30d, hsa-mir-491, hsa-mir-298
MIAT	hsa-mir-29c
NEAT1	hsa-mir-214, hsa-mir-708, hsa-mir-4458, hsa-mir-339, hsa-mir-383, hsa-mir-320a
PVT1	hsa-mir-190b, hsa-mir-488
SNHG1	hsa-mir-4735, hsa-mir-383
SNHG12	hsa-mir-320a
SNHG15	hsa-mir-4735
TUG1	hsa-mir-29c, hsa-mir-31

4.3. mRNAs targeted by miRNAs

To construct the lncRNA-miRNA-mRNA ceRNA regulatory network, mRNA targeted by miRNAs should be searched. miRTarBase and DIANA-LncBase were used to investigate it. For miRTarBase, each miRNA-mRNA pair was experimentally validated by at least three of the following methods, including reporter assay, western blot, qPCR or NGS. For DIANA-LncBase, mRNA targets were predicted via experimental module and had a high Pr. score. Finally, 163 mRNAs were targeted by 21 miRNAs and those with less than 5% genetic alteration screened via cBiPortal were excluded. Finally, 142 mRNAs were predicted as miRNA targets and a total of 162 miRNA-mRNA pairs were obtained (Table 2).

Table 2: Validated mRNAs' targets from TarBase

miRNAs	mRNAs
hsa-mir-135b	ARID1A, MTCH2, STAT6, BGLAP, APC, ACVR1B, EVI5, RUNX2, MAFB, PPP2R5C, SMAD5, MID1, BMPR2, THBS2
hsa-mir-148a	AKT2, USP4, ITGB8, ROCK1, DNMT3B, BCL2L11, CDC25B, CDKN1B, IRS1, SMAD2, TGIF2, ALCAM, CENPF, IKBKB, MAP3K9, RPS6KA5, STAT3, TGFB2, INO80, MAFB, VAV2
hsa-mir-190b	PTEN, BCL2
hsa-mir-214	MAPK14, RAB5B, UBE2I, FLOT1, HDGF, MAPK8, BIRC5, EZH2, KIF1B, TRPS1, BCL2L2, SRGAP1, SRGAP2, ARL2, BCL2L11, CDK3, CDK6, CPD, ING4, MAP2K3, PIM1, RAB15, FGFR1, JAG1, NRAS, PSMD10, RASSF5, TWF1, ARHGAP10, CAPN5, CPEB4, IGF1R, PAPPA, PTEN, QKI, TP53, TWIST1
hsa-mir-296	SCRIB, HGS, WNK4, IKBKE
hsa-mir-298	BACE1, ZFR, MATR3, APP
hsa-mir-29c	AKT2, DNMT3A, CREB5, MCL1, TDG, TFAP2C, DNMT3B, Mmp15, CDK6, RCC2, SIRT1, SRSF10, VEGFA, AKT3, FBN1, GAPDH, BACE1, BCL2, CDC42, COL15A1, CTNND1, PTEN
hsa-mir-30b	RORC, CCNE2, BCL6, SNAI1
hsa-mir-30d	EZH2, RUNX2, CASP3, SNAI1

hsa-mir-31	MET, NUMB, CREG1, RHOA, HOXC13, YY1, NFAT5, SATB2, RET, ARPC5, FZD3, HIF1AN, FOXP3, SELE, RDX, DMD, WASF3
hsa-mir-320a	ARF1, SUZ12, FOXM1, MCL1, RAB14, HOXA10, TRPC5, NFATC3, RUNX2, GNAI1, PTEN
hsa-mir-339	PTP4A1, BCL6, BACE1
hsa-mir-383	ATR, PPP1R10, PRPF31, VEGFA, IGF1R, PRDX3
hsa-miR-4458	NDC1, PXT1, IGF1R
hsa-miR-4735	ZNF507
hsa-mir-488	BCL2L11
hsa-mir-491	EGFR, CAPNS1, GIT1
hsa-mir-708	AKT2, EYA3, BIRC5, MPL
hsa-mir-92b	PRMT5, NLK
hsa-miR-1273a	SMG1
hsa-miR-4756	CCND1

4.4. Construction of a lncRNA-miRNA-mRNA ceRNA network in LUAD

To better understand the role of ncRNAs in LUAD and to further elucidate the interaction between these lncRNAs and miRNAs, a dysregulated lncRNA-miRNA-mRNA ceRNA network of LUAD was constructed. Cytoscape software was used to visualize the network. As shown in Figure 1, a total of 187 nodes and 217 edges were mapped and a total of 24 lncRNAs, 21miRNAs and 142 mRNAs were involved.

Figure 1: The lncRNA-miRNA-mRNA ceRNA network in LUAD. The network was constructed with Attribute Circle Layout algorithm and all lncRNAs and miRNAs were located into the circle.

4.5. Gene ontology (GO) and KEEG pathway analysis

In order to investigate the functional processes involved in LUAD development and progression, mRNAs in the ceRNA network were subjected to Fisher's exact test-based GO biological and functional process analysis. 142 mRNA were submitted to the STRING database and visualized as a PPI network, which consisted of 142 nodes and 765 edges (Figure 2). The results of functional analysis of miRNAs were generated with DAVIAD. miRNAs were categorized into three functional groups: biological process (BP), molecular function (MF) and cellular component (CC). According to number of genes involved, we listed the top 10 significantly related GO analysis as Table 3 shown. The results indicated that in the biological process, up-regulated genes were mainly enriched in negative regulation of transcription from RNA polymerase II promoter, negative regulation of cellular senescence, positive regulation of gene expression, negative regulation of apoptotic pro-

cess and so on. In the cellular component analysis, the genes were mainly enriched in transcription factor complex nucleoplasm, PML body, cytoplasm, nuclear chromatin, focal adhesion, cytosol, nucleus, axon and extracellular exosome. For molecular function, up-regulated genes were significantly enriched in ATP binding, protein serine/threonine kinase activity, RNA polymerase II core promoter sequence-specific DNA binding and so on. The results demonstrated that most mRNAs in the ceRNA network were significantly enriched in biological or cellular process, cell components and kinase activities.

To understand the signal pathways involved in ceRNA network, KEEG pathway analysis was performed. The top 15 KEGG pathways in our study were listed as Table 4 shown. 10 cancer-related pathways were identified, including: MicroRNAs in cancer, Pathways in cancer, Proteoglycans in cancer, Pancreatic cancer, Prostate cancer, Colorectal cancer, Melanoma, Central carbon metabolism in cancer, Glioma and Endometrial cancer. 5 non-cancer related pathways such as Hepatitis B, FoxO signaling pathway, PI3K-Akt signaling pathway, Adherens junction, MAPK signaling pathway were also enriched.
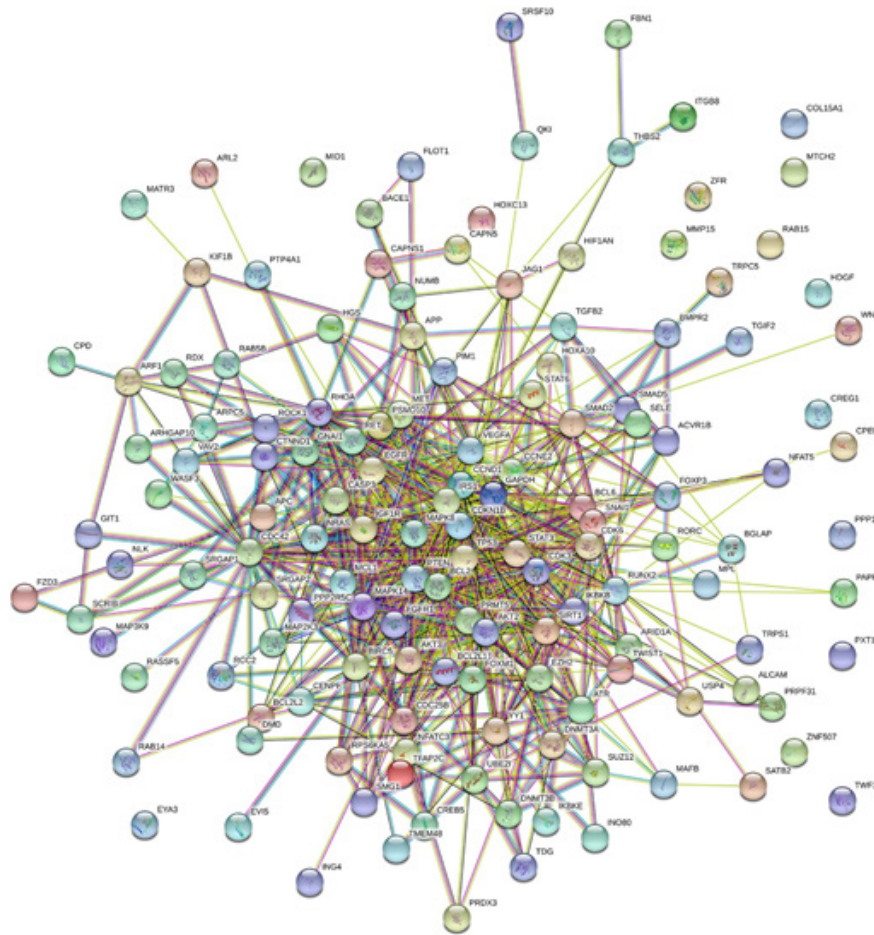
Figure 2: PPI network of mRNAs in the ceRNA network. 142 mRNA were submitted to the STRING database and visualized as a PPI network, which consisted of 142 nodes and 765 edges.

Table 3: Go analysis of mRNAs in the ceRNA network

Term	Count	P Value*
GO-BPs		
GO:0000122~negative regulation of transcription from RNA polymerase II promoter	16	1.27E-06
GO:2000773~negative regulation of cellular senescence	4	4.04E-05
GO:0010628~positive regulation of gene expression	9	4.96E-05
GO:0043066~negative regulation of apoptotic process	11	7.82E-05
GO:0045944~positive regulation of transcription from RNA polymerase II promoter	15	2.72E-04
GO:0032000~positive regulation of fatty acid beta-oxidation	3	0.00111
GO:0071479~cellular response to ionizing radiation	4	0.002047
GO:0006978~DNA damage response, signal transduction by p53 class mediator resulting in transcription of p21 class mediator	3	0.002299
GO:0001657~ureteric bud development	4	0.002308

GO:0007283~spermatogenesis	7	0.004424
GO-CCs		
GO:0005667~transcription factor complex	8	7.24E-05
GO:0005654~nucleoplasm	25	1.34E-04
GO:0016605~PML body	6	1.35E-04
GO:0005737~cytoplasm	34	2.10E-04
GO:0000790~nuclear chromatin	7	3.74E-04
GO:0005925~focal adhesion	10	9.86E-04
GO:0005829~cytosol	14	0.007474
GO:0005634~nucleus	27	0.008162
GO:0030424~axon	5	0.008597
GO:0070062~extracellular exosome	27	0.031307
GO-MFs		
GO:0005524~ATP binding	27	7.72E-05
GO:0004674~protein serine/threonine kinase activity	10	1.33E-04
GO:0000979~RNA polymerase II core promoter sequence-specific DNA binding	4	1.13E-02
GO:0004713~protein tyrosine kinase activity	3	1.98E-02
GO:0001077~transcriptional activator activity, RNA polymerase II core promoter proximal region sequence-specific binding	6	2.61E-02
GO:0003700~transcription factor activity, sequence-specific DNA binding	8	3.08E-02
GO:0008384~IkappaB kinase activity	2	0.030858
GO:0003886~DNA (cytosine-5-)-methyltransferase activity	2	0.030858
GO:0005525~GTP binding	8	0.038338
GO:0008201~heparin binding	4	0.043016

*BP, biological process; MF, molecular function; CC, cellular component. P value < 0.05 was considered as threshold values of significant difference.

Table 4: KEGG pathway analysis of mRNAs in the ceRNA network

KEEG pathways	**Count**	**P Value**[*]
Cancer related		
oas05206: MicroRNAs in cancer	31	3.44E-27
oas05200: Pathways in cancer	28	1.40E-14
oas05205: Proteoglycans in cancer	19	1.27E-11
oas05212: Pancreatic cancer	12	1.54E-10
oas05215: Prostate cancer	13	2.13E-10
oas05210: Colorectal cancer	11	1.82E-09
oas05218: Melanoma	10	9.46E-08
oas05230: Central carbon metabolism in cancer	9	5.69E-07
oas05214: Glioma	8	8.33E-06
oas05213: Endometrial cancer	7	2.56E-05
Non-cancer related		
oas05161: Hepatitis B	19	2.59E-14
oas04068: FoxO signaling pathway	18	1.02E-13
oas04151: PI3K-Akt signaling pathway	22	2.62E-10
oas04520: Adherens junction	11	5.30E-09
oas04010: MAPK signaling pathway	17	2.24E-08

P value < 0.05 was considered as threshold values of significant difference.

4.4. Potential prognosis of LUAD specific lncRNAs in the ceRNA network

To identify the potential lncRNAs with prognostic characteristics, the expression levels of 24 LUAD specific lncRNAs in the ceRNA network were profiled using survival Kaplan-Meier estimate from cBioPortal. As a result, using Onco Query Language (OQL) EXP>2, three lncRNAs (CASC15, LINC00696 and LINC01600) were found to be significantly associated with overall survival (P < 0.05) (Figure 3A-C). The combination of these three lncRNAs

accounted for a total of 17.5% (90/ 517) cases and negatively correlated with OS (P = 8.130E-6) (Figure 3D). Recurrence is a major concern for cancer survivors. There were 187 cases of recurrence in this cohort. The results indicated that four lncRNAs (CASC15, LINC00662, LINC01600 and MALAT1) was significantly associated with recurrence (p <0.05). Of those, three lncRNA, including CASC15, LINC00662 and LINC01600 were negatively correlated with progression-free survival (P < 0.05) (Figure 3E-G). On the contrary, MALAT1 were positively correlated with progression-free survival time (P< 0.05) (Figure 3H).
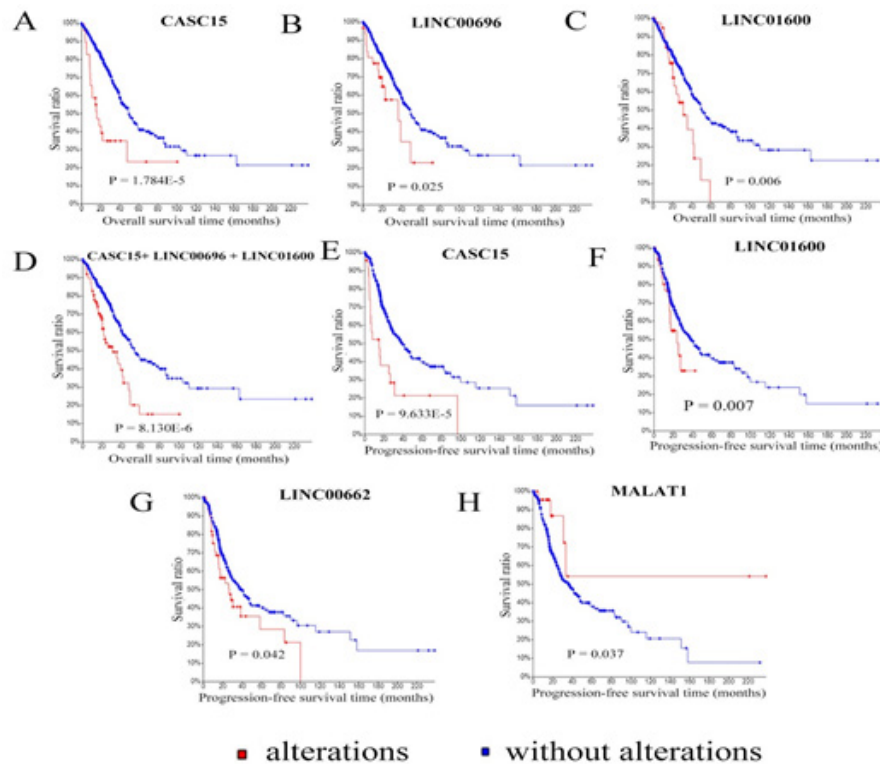
Figure 3: Kaplan–Meier analysis of prognosis related lncRNAs. (A-C) Kaplan–Meier curves of CASC15, LINC00696 and LINC01600 were associated with OS. (D) Kaplan–Meier curve of the combination of the above three lncRNAs. (E–H) Kaplan–Meier curves of CASC15, LINC00662, LINC01600 and MALAT1 were associated with recurrence. Red and blue dots indicate patient samples with or without genetic alterations, respectively. P-value indicates the significance of difference.

4.5. Clinical significance of CASC15 and LINC01600 in LUAD

Based on above results, we noticed that two lncRNAs, CASC15 and LINC01600, performed the potential prognostic characteristics with LUAD regarding both OS and recurrence, suggesting they may exist as a synergic relationship. The combination of CASC15 and LINC01600 dramatically negatively correlated with OS and recurrence (P <0.05) (Figure 4A and 4B). To reveal the clinical significance of these two lncRNAs in LUAD, genetic alterations, as well as alterations to other clinical features, such as sex, age in diagnosis, tumor stage, metastatic status, lymph node metastasis and KRAS mutation, were investigated. Notably, LINC01600 was identified to be associated with KRAS mutation. As presented in Figure 4C, the expression of LINC01600 in cases without KRAS mutation was significantly increased compared with cases with KRAS mutation (P<0.05). Furthermore, we found that the

expression of LINC01600 in cases with metastasis (M =1 or x) was significantly increased compared with cases without metastasis (M = 0) (P<0.05) (Figure 4E). However, there were no associations between CASC15 and metastasis status or KRAS mutation (Figure 4D and 4F). Other clinical features, including sex, age in diagnosis, tumor stage and lymph node metastasis, were evaluated separately, and no significant associations with CASC15 or LINC01600 were observed.

CBioPortal provided the probable co-occurrence of these two lncRNAs and there was a significant tendency towards co-occurrence between CASC15 and LINC01600 in LUSC (P = 0.021). The co-expressed genes of these two lncRNAs were also determined. As a result, 31 genes were revealed to be co-expressed with CASC15, and 125 genes were discovered to have a co-expressed relationship with LINC01600, based on the algorithm blow: Pear-

son's Correlation > 0.5 (Supplementary Table 5). Go analysis indicated that biological and functional process of CASC15 and LINC01600 was similar. In the biological process, both CASC15 and LINC01600 co-expressed genes were most enriched in extra-

cellular exosome or space and in the cellular component analysis, both of them were most significantly involved in interferon-gamma-mediated signaling pathway, signal peptide processing and regulation of protein stability.
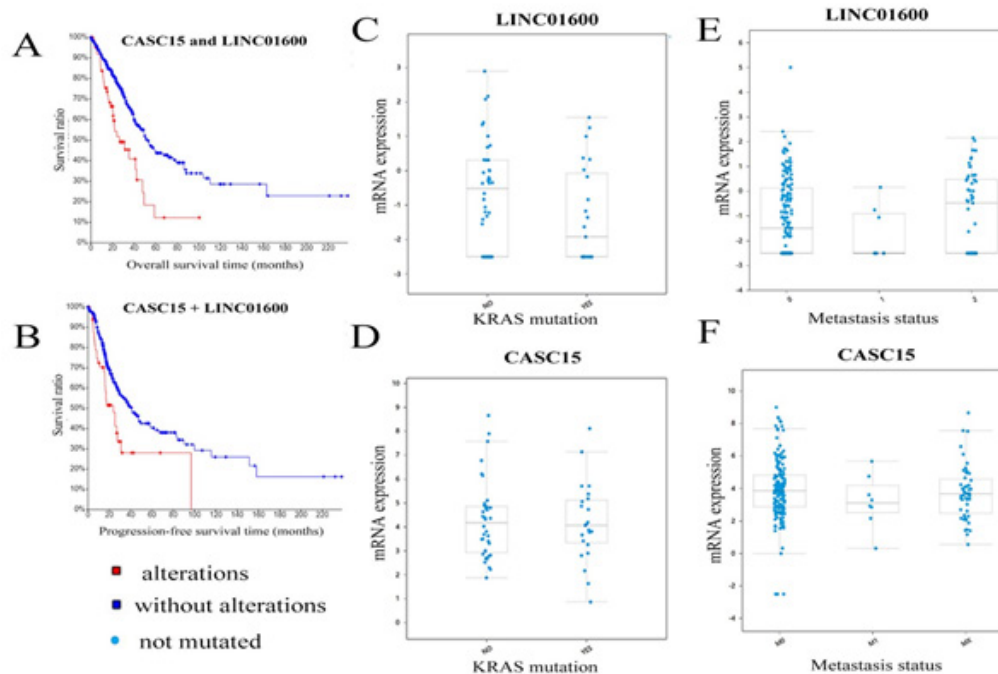

Figure 4: The prognostic characteristics of CASC15 and LINC01600 regarding OS, recurrence and other clinical features (lymph node metastasis and KRAS mutation). (A-B) The combination of CASC15 and LINC01600 dramatically negatively correlated with OS and recurrence (P <0.05); (C-F) the expression of LINC01600 in cases with metastasis or KRAS mutation was significantly increased compared with cases without metastasis or KRAS mutation. However, there were no associations between CASC15 and metastasis status or KRAS mutation.

5. Discussion

Recently, aberrant expression of lncRNAs has been widely observed in various cancers and accumulative evidence indicates that lncRNAs emerge as key regulatory roles in tumorigenesis and tumor progression [27-30]. Notably, in addition to directly interacted with mRNAs, lncRNAs can also function as endogenous miRNA sponges of a lncRNA-miRNA-mRNA based ceRNA network to indirectly regulate targeted mRNAs and several vital regulatory pathways, as well as potential therapeutic targets, may be revealed. Wang et al. found that CircNT5E acted as a sponge of microRNA-422a to promote glioblastoma tumorigenesis [31]. Mou et al. described that lncRNA-ATB functioned as a competing endogenous RNA to promote YAP1 by sponging miR-590-5p in malignant melanoma [32]. The above results elucidated the important roles of lncRNAs in tumorigenesis and tumor development as a part of ceRNA networks and suggested a potential strategy for the treatment of cancers.

To date, some efforts had been made to describe lncRNA profiles in LUAD and several dysregulated lncRNAs as ceRNAs were reported based on microarray analysis. Li et al. revealed that MEG3 and MIAT regulate MAPK9 by interactions with miR-106 to involve in MAPK signaling pathways and LINC00115 might inter-

act with miR-7 to regulate FGF2 to participate in pathways in cancer [33]. These lncRNAs may be underlying therapeutic targets for LUAD functioning as ceRNAs for regulation of miRNA-mRNA. Sui et al. found 21 lncRNAs in LUAD-related ceRNA network were aberrantly expressed with clinical features and 5 lncRNAs (BCRP3, LINC00472, CHIAP2, BMS1P20 and UNQ6494) positively correlated with overall survival (OS) [34]. Zhao et al. constructed a specific SVM (support vector machine) classifier based on the ceRNA network for diagnosis of LUAD and identified hsa-miR-96, hsa-miR-204, PGM5P2, SFTA1P, RGS20, RGS9BP, FGB and INA may serve as prognostic markers in clinical practice [26]. However, studies on identification of the LUAD-related lncRNAs based on genome-wide RNA profiles and large sample size were still poorly described.

In the current study, we interrogated the LUAD dataset of TCGA with larger 517 tissue samples at cBioPortal with the latest updated lncRNAs and miRNAs database from HGNC. lncRNAs and miRNAs with highly genetic alterations were identified and subsequently selected to construct the regulated lncRNA-miRNA-mRNA based ceRNA networks. A total of 24 lncRNAs, 19 miRNAs and 142 mRNAs were involved in this network and GO and KEGG pathway analysis of targeted mRNAs was operated. The results of

GO analysis indicated that most mRNAs in the ceRNA network were significantly enriched in biological or cellular process, cell components and kinase activities. Based on KEGG pathway analysis, in addition cancer-related pathways, some non-cancer related pathways, including FoxO signaling pathway, PI3K-Akt signaling pathway and MAPK signaling pathway were also enriched. Some studies had reported that these signaling pathways were involved in the initiation and progression of LUAD [35]. Furthermore, among 24 lncRNAs in the network, several key lncRNAs were identified as LUAD prognosis-related. Three lncRNAs (CASC15, LINC00696 and LINC01600) were found to be negatively correlated with OS and four lncRNAs (CASC15, LINC00662, LINC01600 and MALAT1) was significantly associated with recurrence. Of those, CASC15, LINC00662 and LINC01600 were negatively correlated with recurrence, and conversely, MALAT1 were positively correlated with recurrence. LINC00696, LINC01600 and LINC00662 were novel identified LUAD prognosis-related lncRNAs and their functions remained elusive. MALAT1 was a well-known lncRNA, which can function as a ceRNA by sponging to miRNAs to indirectly regulate targeted mRNAs in various cancers. Luan et al. found that MALAT1 acted as a ceRNA to promote malignant melanoma growth and metastasis by sponging miR-22 [36]. Zhang et al revealed that MALAT1 regulated the expression of Gli2 by miR-202 to strengthen gastric cancer progression [37]. Tao et al discovered that miR-211 sponged MALAT1 to suppress tumor growth and progression through inhibiting PHF19 in ovarian carcinoma [38]. Chang et al. found that MALAT1 can function as a ceRNA to modulate STAT3 expression by absorbing miR-125b in oral squamous cell carcinoma (OSCC) and could be used as a novel therapeutic target in OSCC diagnosis and treatment [39]. CASC15, cancer susceptibility candidate 15, originally named FLJ22536, was a long intergenic non-coding RNA (lincRNA) locus in chromosome 6p22.3 [40]. Lessard et al. reported that CASC15 was involved in melanoma progression and phenotype switching, and was an independent predictor of recurrence for metastatic melanoma [41]. Fernando et al found that CASC15 can regulate SOX4 expression in RUNX1-rearranged acute leukemia [42]. Some other researches described overexpression of CASC15 may promote tumor development and progression in gastric cancer and hepatocellular carcinoma, and was correlated with a poor prognosis [43,44]. However, studies about how CASC15 functioned as a ceRNA were rare. Recently, Jing et al. carried out a study and the results revealed that CASC15 promoted colon cancer growth and metastasis through the activation of the Wnt/β-catenin signaling pathway by acting as a sponge to suppress miR-4310 that targeted LGR5 [45]. This study suggested that CASC15 may be a therapeutic target for colon cancer treatment.

After systemic analysis of LUAD prognosis-related lncRNAs, interestingly, we found that CASC15 and LINC01600 both had the potential prognostic characteristics with LUAD regarding OS and recurrence, which drove us to explore their relationships and the correlation between them and other clinical features. We found that there were significant associations between LINC01600 and KRAS mutation or metastasis status, suggesting LINC01600 may be an oncogene and play a role in the KRAS related signaling pathways of LUAD. However, the associations between CASC15 and LINC01600 and other clinical features, including sex, age in diagnosis, tumor stage and lymph node metastasis, were observed. Moreover, bioinformatics analysis indicated that there was a significant tendency towards co-occurrence between CASC15 and LINC01600 in LUSC (P = 0.001). The results of Go analysis showed that the co-expressed genes of CASC15 and LINC01600 were similar, which were both enriched in extracellular exosome or space, and most significantly involved in interferon-gamma-mediated signaling pathway, signal peptide processing and regulation of protein stability. The similarity of co-expressed genes of CASC15 and LINC01600 may partly explain the reason why they had a significant tendency towards co-occurrence.

In this study, several notes should be mentioned. Firstly, for increasing the accuracy of ceRNA network prediction, we only recruited the cancer specific lncRNAs and miRNAs that had alterations more than 5% and were annotated by HGNC. The relationships between lncRNA and miRNA, and miRNA and mRNA were predicted by experiment-supported algorithms or databases such as starBase and miRTarBase. These measurements guaranteed that the relationships identified would be reliable not only in silico situations but also by experimental-supported evidences. However, the limitation of the constructed ceRNA network and LUAD prognosis-related lncRNAs were still obvious due to the lack of experiment validation. Further experiments, such as reporter assay, western blot, or qPCR methods were mandatory to confirm these results. Moreover, since TCGA was a public and accessible cancer database, it was inevitable to come across the situation that results were obtained from different studies based on the same database. For example, Zhao also used tumor sample data from TCGA to construct a LUAD related miRNA-lncRNA-mRNA network and identified molecular markers with diagnostic and prognostic value for LUAD26. However, there were still some differences between us. Firstly, the numbers and names of molecules in the constructed ceRNA network were different: there were 6 lncRNAs, 25 miRNAs and 126 mRNAs in Zhao's work; in our study, the numbers of lncRNAs, miRNAs and mRNAs were 24, 21 and 142, respectively. Secondly, 2 lncRNAs, 2 miRNAs and 4 mRNAs were identified to be prognosis related in Zhao's work while we focused on lncRNAs. Molecular markers in Zhao's study were associated with OS; however, we found several lncRNAs performed the potential prognostic characteristics with LUAD regarding not only OS, but also recurrence, which revealed their relationships and the correlation between them and other clinical features.

In conclusion, in the present study, we draw aberrant expression

profiles of LUAD-related lncRNAs and miRNAs from hundreds of candidate lncRNAs detected from large samples size in TCGA database and constructed the lncRNA-miRNA-mRNA ceRNA network to clarify the unknown ceRNA regulatory network in LUAD. Some key lncRNAs were subsequently identified as LUAD prognosis-related, and of those, LINC01600 and CASC15 both performed the potential prognostic characteristics with LUAD regarding OS and recurrence. Comprehensive analysis indicated that the expression of LINC01600 was significantly associated with KRAS mutation and lymph node metastasis, and CASC15 and LINC01600 were significantly tended towards co-occurrence, which may be due to the similarity of genes co-expressed by these two lncRNAs. Our findings provided novel insight into better understanding of the lncRNA-miRNA-mRNA based ceRNA network in LUAD and potential biomarkers for prognosis.

6. Funding

Reference

1. Wild CP, Weiderpass E, Stewart BW. World Cancer Report: Cancer Research for Cancer Prevention. Lyon, France: International Agency for Research on Cancer. 2020.

2. Cancer Genome Atlas Research, N. Comprehensive molecular profiling of lung adenocarcinoma. Nature. 2014; 511: 543-550.

3. Planck M, et al. Genomic and transcriptional alterations in lung adenocarcinoma in relation to EGFR and KRAS mutation status. PLoS One. 2013; 8: e78614.

4. Takahashi T, et al. p53: a frequent target for genetic abnormalities in lung cancer. Science. 1989; 246: 491-494.

5. Devarakonda S, Morgensztern D, Govindan R. Genomic alterations in lung adenocarcinoma. Lancet Oncol. 2015; 16: e342-351.

6. Morris KV, Mattick JS. The rise of regulatory RNA. Nature reviews. Genetics. 15: 423-437.

7. Liz J, Esteller M. lncRNAs and microRNAs with a role in cancer development. Biochim Biophys Acta. 2016; 1859; 169-176.

8. Garzon R, Calin GA, Croce CM. MicroRNAs in Cancer. Annu Rev Med. 2009; 60: 167-179.

9. Cantini L, et al. MicroRNA-mRNA interactions underlying colorectal cancer molecular subtypes. Nat Commun. 2015; 6: 8878.

10. Ooi CY, et al. Network Modeling of microRNA-mRNA Interactions in Neuroblastoma Tumorigenesis Identifies miR-204 as a Direct Inhibitor of MYCN. Cancer Res. 2018; 78: 3122-3134.

11. Huarte M. The emerging role of lncRNAs in cancer. Nat Med. 2015; 21: 1253-1261.

12. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? Cell. 2011; 146; 353-358.

13. Ebert MS, Neilson JR, Sharp PA. MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. Nat Methods. 2007; 4: 721-726.

14. Poliseno L, et al. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. Nature. 2010; 465: 1033-1038.

15. Tay Y, Rinn J, Pandolfi PP. The multilayered complexity of ceRNA crosstalk and competition. Nature. 2014; 505: 344-352.

16. Tay Y, et al. Coding-independent regulation of the tumor suppressor PTEN by competing endogenous mRNAs. Cell. 2011; 147: 344-357.

17. Kallen AN, et al. The imprinted H19 lncRNA antagonizes let-7 microRNAs. Mol Cell. 2013; 52: 101-12.

18. Li JH, Liu S, Zhou H, Qu LH, Yang JH. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. Nucleic Acids Res. 2014; 42: D92-97.

19. Chou CH, et al. miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. Nucleic Acids Res. 2018; 46: D296-D302.

20. Karagkouni D, et al. DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA-gene interactions. Nucleic Acids Res. 2018; 46: D239-D245.

21. Paraskevopoulou MD, et al. DIANA-LncBase v2: indexing microRNA targets on non-coding transcripts. Nucleic Acids Res. 2016; 44: D231-238.

22. Xia T, et al. Long noncoding RNA associated-competing endogenous RNAs in gastric cancer. Sci Rep. 2014; 4: 6088.

23. Zhou X, Liu J, Wang W. Construction and investigation of breast-cancer-specific ceRNA network based on the mRNA and miRNA expression data. IET Syst Biol. 2014; 8: 96-103.

24. Wang H, et al. Comprehensive analysis of aberrantly expressed profiles of lncRNAs and miRNAs with associated ceRNA network in muscle-invasive bladder cancer. Oncotarget. 2016; 7: 86174-86185.

25. Sui J, et al. Comprehensive analysis of aberrantly expressed microRNA profiles reveals potential biomarkers of human lung adenocarcinoma progression. Oncol Rep. 2017; 38: 2453-2463.

26. Zhao J, et al. Construction of a specific SVM classifier and identification of molecular markers for lung adenocarcinoma based on lncRNA-miRNA-mRNA network. Onco Targets Ther. 2018; 11: 3129-3140.

27. Martens-Uzunova ES, et al. Long noncoding RNA in prostate, bladder, and kidney cancer. Eur Urol 2014; 65: 1140-1151.

28. Zhang J, et al. A novel TGF-beta and H19 signaling axis in tumor-initiating hepatocytes that regulates hepatic carcinogenesis. Hepatology. 2018.

29. Zhang X, et al. Long noncoding RNA HNF1A-AS1 indicates a poor prognosis of colorectal cancer and promotes carcinogenesis via activation of the Wnt/beta-catenin signaling pathway. Biomed Pharmacother. 2017; 96: 877-883.

30. Zhang D, Sun G, Zhang H, Tian J, Li Y. Long non-coding RNA AN-RIL indicates a poor prognosis of cervical cancer and promotes carcinogenesis via PI3K/Akt pathways. Biomed Pharmacother. 2017; 85: 511-516.

31. Wang R, et al. CircNT5E acts as a sponge of microRNA-422a to promote glioblastoma tumorigenesis. Cancer Res. 2018.

32. Mou K, et al. lncRNA-ATB functions as a competing endogenous RNA to promote YAP1 by sponging miR-590-5p in malignant melanoma. Int J Oncol. 2018.

33. Li DS, Ainiwaer JL, Sheyhiding I, Zhang Z, Zhang LW. Identification of key long non-coding RNAs as competing endogenous RNAs for miRNA-mRNA in lung adenocarcinoma. Eur Rev Med Pharmacol Sci. 2016; 20: 2285-2295.

34. Sui J, et al. Integrated analysis of long non-coding RNAassociated ceRNA network reveals potential lncRNA biomarkers in human lung adenocarcinoma. Int J Oncol. 2016; 49: 2023-2036.

35. Wang J, et al. BMP9 inhibits the growth and migration of lung adenocarcinoma A549 cells in a bone marrow stromal cellderived microenvironment through the MAPK/ERK and NF-kappaB pathways. Oncology reports. 2016; 36; 410-418.

36. Luan W, et al. Long non-coding RNA MALAT1 acts as a competing endogenous RNA to promote malignant melanoma growth and metastasis by sponging miR-22. Oncotarget. 2016; 7: 63901-63912.

37. Zhang Y, Chen Z, Li MJ, Guo HY, Jing NC. Long non-coding RNA metastasis-associated lung adenocarcinoma transcript 1 regulates the expression of Gli2 by miR-202 to strengthen gastric cancer progression. Biomed Pharmacother. 2017; 85: 264-271.

38. Tao F, Tian X, Ruan S, Shen M, Zhang Z. miR-211 sponges lncRNA MALAT1 to suppress tumor growth and progression through inhibiting PHF19 in ovarian carcinoma. FASEB J. 2018; fj201800495RR.

39. Chang SM, Hu WW. Long non-coding RNA MALAT1 promotes oral squamous cell carcinoma development via microRNA-125b/STAT3 axis. J Cell Physiol. 2018; 233: 3384-3396.

40. Maris JM, et al. Chromosome 6p22 locus associated with clinically aggressive neuroblastoma. The New England journal of medicine. 2008; 358, 2585-2593.

41. Lessard L, et al. The CASC15 Long Intergenic Noncoding RNA Locus Is Involved in Melanoma Progression and Phenotype Switching. The Journal of investigative dermatology. 2015; 135: 2464-2474.

42. Fernando TR, et al. The lncRNA CASC15 regulates SOX4 expression in RUNX1-rearranged acute leukemia. Molecular cancer. 2017; 16: 126.

43. Yao XM, Tang JH, Zhu H, Jing Y. High expression of LncRNA CASC15 is a risk factor for gastric cancer prognosis and promote the proliferation of gastric cancer. European review for medical and pharmacological sciences. 2017; 21: 5661-5667.

44. He T, et al. Long non-coding RNA CASC15 is upregulated in hepatocellular carcinoma and facilitates hepatocarcinogenesis. International journal of oncology. 2017; 51: 1722-1730.

45. Jing N, et al. LncRNA CASC15 promotes colon cancer cell proliferation and metastasis by regulating the miR4310/LGR5/Wnt/betacatenin signaling pathway. Molecular medicine reports. 2018; 18: 2269-2276.